

CREATION AND GROWTH OF COMPONENTS IN A RANDOM HYPERGRAPH PROCESS

VLADY RAVELOMANANA AND ALPHONSE LAZA RIJAMAMY

ABSTRACT. Denote by an ℓ -component a connected b -uniform hypergraph with k edges and $k(b-1) - \ell$ vertices. We prove that the expected number of creations of ℓ -component during a random hypergraph process tends to 1 as ℓ and b tend to ∞ with the total number of vertices n such that $\ell = o(\sqrt[3]{\frac{n}{b}})$. Under the same conditions, we also show that the expected number of vertices that ever belong to an ℓ -component is approximately $12^{1/3}(b-1)^{1/3}\ell^{1/3}n^{2/3}$. As an immediate consequence, it follows that with high probability the largest ℓ -component during the process is of size $O((b-1)^{1/3}\ell^{1/3}n^{2/3})$. Our results give insight about the size of giant components inside the phase transition of random hypergraphs.

1. INTRODUCTION

A *hypergraph* \mathcal{H} is a pair $(\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, 2, \dots, n\}$ denotes the set of vertices of \mathcal{H} and \mathcal{E} is a family of subsets of \mathcal{V} called edges (or hyperedges). For a general treatise on hypergraphs, we refer to Berge [2]. We say that \mathcal{H} is *b-uniform* (or simply *uniform*) if for every edge $e \in \mathcal{E}$, $|e| = b$. In this paper, all considered hypergraphs are b -uniform. We will study the growth of size and complexity of connected components of a random hypergraph process $\{\mathbb{H}(n, t)\}_{0 \leq t \leq 1}$ defined as follows. Let K_n be the complete hypergraph built with n vertices and $\binom{n}{b}$ edges (self-loops and multiple edges are not allowed). $\{\mathbb{H}(n, t)\}_{0 \leq t \leq 1}$ may be constructed by letting each edge e of K_n (amongst the $\binom{n}{b}$ possible edges) appear at random time T_e , with T_e independent and uniformly distributed on $(0, 1)$ and letting $\{\mathbb{H}(n, t)\}_{0 \leq t \leq 1}$ contain the edges such that $T_e \leq t$ (for the random graph counterpart of this model, we refer the reader to [17, 24]). This model is closely related to $\{\mathbb{H}(n, M)\}$ where $M \in [1, \binom{n}{b}]$ represents the number of edges picked uniformly at random amongst the $\binom{n}{b}$ possible edges and which are present in the random hypergraph. The main difference between $\{\mathbb{H}(n, M)\}_{0 \leq M \leq \binom{n}{b}}$ and $\{\mathbb{H}(n, t)\}_{0 \leq t \leq 1}$ is that in $\{\mathbb{H}(n, M)\}_{0 \leq M \leq \binom{n}{b}}$, edges are added at fixed (slotted) times $1, 2, \dots, \binom{n}{b}$ so at any time M we obtain a random graph with n vertices and M edges, whereas in $\{\mathbb{H}(n, t)\}_{0 \leq t \leq 1}$ the edges are added at random times. At time $t = 0$, we have a hypergraph with n vertices and 0 edge, and as the time advances all edges e with r.v. T_e such that $T_e \leq t$ (where t is the current time), are added to the hypergraph until t reaches 1 in which case, one obtains the complete hypergraph K_n .

We define the *excess* (or the *complexity*) of a connected b -uniform hypergraph as (see also [20]):

$$(1) \quad \text{excess}(\mathcal{H}) = \sum_{e \in \mathcal{E}} (|e| - 1) - |\mathcal{V}| = |\mathcal{E}| \times (b - 1) - |\mathcal{V}|.$$

Key words and phrases. Random hypergraphs; probabilistic/analytic combinatorics; asymptotic enumeration; extremal hypergraphs.

Namely, the complexity (or excess) of connected components ranges from -1 (hypertrees) to $\binom{n}{b}(b-1) - n$ (complete hypergraph). A connected component with excess ℓ ($\ell \geq -1$) is called an ℓ -*component*. The notion of excess was first used in [29] where the author obtained substantial enumerative results in the study of connected graphs according to the two parameters number of vertices and number of edges. It was also used in enumerative combinatorics and as well as in various study of random hypergraphs processes [20, 21, 1].

Numerous results have been obtained for random graphs as witnessed by the books [4, 19] and the references therein. In comparison, there are very few works about random hypergraphs. One of the most significant results was obtained by Schmidt-Prznan and Shamir [26] who studied the component structure for random hypergraphs. In particular, they proved that if $b \geq 2$, $M = cn$ with $c < 1/b(b-1)$ then asymptotically almost surely (a.a.s. for short) the largest component of $\mathbb{H}(n, M)$ is of order $\log n$ and for $c = 1/b(b-1)$ it has $\theta(n^{2/3})$ vertices and as $c > 1/b(b-1)$ a.a.s. $\mathbb{H}(n, M)$ has a unique geant component with $\theta(n)$ vertices. This result generalizes the seminal papers of Erdős and Rényi who discovered the abrupt change in the structure of the random graph $\mathbb{G}(n, M)$ when $M = cn$ with $c \sim 1/2$. In [21], Karoński and Łuczak proved limit theorems for the distribution of the size of the largest component of $\mathbb{H}(n, M)$ at the phase transition, i.e., $M = n/b(b-1) + O(n^{2/3})$.

In this paper, we consider the *continuous time* random hypergraph process described above and will study the creation and growth of components of excess ℓ (or ℓ -components). A connected component which is not a hypertree is said *multicyclic* (following the terms used by our predecessors in [16, 17, 18]).

1.1. Definitions. We can observe that there are two manners to create a new $(\ell + 1)$ component during the $\{\mathbb{H}(n, t)\}_{0 \leq t \leq 1}$ process :

- either by adding an edge between an existing p -component (with $p \leq \ell$) and $(b - q)$ hypertrees (with $0 \leq q \leq b$) such that the edge encloses q distinct vertices in the p -component,
- or by joining with the last added edge many connected components such that the number of multicyclic components diminishes.

Observe that in the first case, to create an ℓ -component, we must have $p + q - 1 = \ell$. In this case, it is also important to note that the number of multicyclic components remains the same after the addition of the last edge.

The first transition described above will be denoted $p \rightarrow \ell$ and the second $\oplus_i p_i \rightarrow \ell$. We say that an ℓ -component is *created* by a transition $p \rightarrow \ell$ with $p < \ell$ or by a transition $\oplus_i p_i \rightarrow \ell$. For $\ell \geq 0$, we say that an ℓ -component *grows* when it swallows some hypertrees (transition $\ell \rightarrow \ell$).

Following Janson in [17], we have two points of view :

- *The static view.* Let $\mathcal{C}_\ell(m)$ denote the collection of all ℓ -components in $\{\mathbb{H}(n, t)\}_{0 \leq t \leq 1}$. Consider the family $\mathcal{C}_\ell^* = \bigcup_m \mathcal{C}_\ell(m)$ for every ℓ -component that appears at some stage of the continuous process, ignoring when it appears : the elements of \mathcal{C}_ℓ^* are called *static ℓ -components*.
- *The dynamic view.* A connected component can be viewed as “the same” according to its excess even after it has grown by swallowing some hypertrees (transition $\ell \rightarrow \ell$). Such component whose excess remains the same can be viewed as a *dynamic ℓ -component* as its size evolves.

We define $V_\ell = |\mathcal{V}_\ell|$ as the number of vertices that at some stage of the process belong to an ℓ -component and $V_\ell^{max} = \max\{|V(C)| : C \in \mathcal{C}_\ell^*\}$ to be the size of the largest ℓ -component that ever appears.

Let $\alpha(\ell; k)$ be the expected number of times a new edge is added by means of the first type of transition $p \rightarrow \ell$ in order to create an ℓ -component with k edges (or with $k \times (b - 1) - \ell$ vertices). Note again that in this case, the number of multicyclic components of the $\{\mathbb{H}(n, t)\}_{0 \leq t \leq 1}$ process remains the same after the addition of this edge.

Similarly, let $\beta(\ell; k)$ be the expected number of times an edge is added joining at least two multicyclic components in order to form a newly ℓ -component with a total of k edges. In other terms, $\beta(\ell; k)$ is the expected number of times at least two multicyclic components and some hypertrees merge to form an ℓ -component.

1.2. Our results and outline of the paper. We combine analytic combinatorics [11] and probabilistic theory [19] to study the extremal characteristics of the components of a random hypergraph process inside its phase transition [21] and find that the size of the largest component with k (hyper)edges and $k(b - 1) - \ell$ vertices is of order $O((b - 1)^{1/3} \ell^{1/3} n^{2/3})$.

This extended abstract is organized as follows. In the next section, we introduce the general expression of the expectations of several random variables of our interest. In section 3, the computations of the expectations are developed focusing on the particular and instructive case of unicyclic components. The last paragraph provides several technical lemmas useful in order to study the extremal case, i.e. whenever the excess ℓ of the component is large. We give there methods on how to investigate the number of creations of ℓ -components as well as the expectation of their size.

2. CONNECTED COMPONENTS AND EXPECTATION OF TRANSITIONS

2.1. Expected number of transitions. In this paragraph, we give a general formal expression of the expectation of the number of the first (resp. second) type of transitions $\alpha(\ell; k)$ (resp. $\beta(\ell; k)$).

We have the following lemma which computes the expected number of transitions $\alpha(\ell; k)$:

Lemma 2.1. *Let $a = k(b - 1) - \ell$. Denote by $\rho(a, k)$ the number of manners to label an ℓ -component with a vertices such that one edge – whose deletion will not increase the number of multicyclic components but will suppress the newly created ℓ -component – is distinguished among the others. Then,*

$$(2) \quad \alpha(\ell; k) = \binom{n}{a} \rho(a, k) \int_0^1 t^{k-1} (1-t)^{\binom{n}{b} - \binom{n-a}{b} - k} dt.$$

Proof. There are $\binom{n}{a}$ choices of the $a = k(b - 1) - \ell$ vertices of the newly created ℓ -component. By the definition of $\rho(a, k)$, there are $\binom{n}{a} \times \rho(a, k)$ possible ℓ -components. The probability that the previous component (the one before obtaining the current ℓ -component) belongs to $\{\mathbb{H}(n, t)\}_{0 \leq t \leq 1}$ is given by

$$(3) \quad t^{k-1} (1-t)^{\sum_{i=1}^{b-1} \binom{n-a}{i} \binom{a}{b-i} + \binom{a}{b} - k + 1}$$

where the summation in the exponent represents the number of edges not present between the considered component and the rest of the hypergraph. The conditional probability that

the last edge is added during the time interval $(t, t + dt)$ and not earlier is $dt/(1 - t)$. Using the identity

$$(4) \quad \sum_{i=1}^{b-1} \binom{n-a}{i} \binom{a}{b-i} = \binom{n}{b} - \binom{n-a}{b} - \binom{a}{b}$$

and integrating over all times after some algebra, we obtain (2). \square

Similarly, if we let $\tau(a; k)$ to be the number of ways to label an ℓ -component with $a = (k-1) - \ell$ vertices and k edges such that one edge – whose suppression augments the number of multicyclic connected components – is distinguished among the others. Then, $\beta(\ell; k)$ can be computed as for $\alpha(\ell; k)$ using exactly $\tau(a; k)$ instead of $\rho(a; k)$.

Next, the following lemma gives some asymptotic values needed when using formula (2).

Lemma 2.2. *Let $a = (b-1)k - \ell$. We have*

$$(5) \quad \begin{aligned} & \binom{n}{a} \int_0^1 t^{k-1} (1-t)^{\binom{n}{b} - \binom{n-a}{b} - k} dt = \frac{1}{\sqrt{(b-1)n^\ell}} \frac{k^{(k-1)} [(b-1)!]^k}{\left(k(b-1) - \ell\right)^{kb-\ell}} \\ & \times \exp\left(k(b-2) - \ell - \frac{(b-1)^4 k^3}{24n^2}\right) \times \left(1 + O\left(\frac{bk}{n} + \frac{b^4 k^2}{n^2} + \frac{b^4 k}{n^2} + \frac{b^4 k^4}{n^3} + \frac{k}{n^{b-1}b} + \frac{1}{k}\right)\right). \end{aligned}$$

Proof. First, using Stirling formula for factorial we get

$$(6) \quad \binom{n}{a} = \frac{1}{\sqrt{2\pi a}} \frac{n^a e^a}{a^a} \exp\left(-\frac{a^2}{2n} - \frac{a^3}{6n^2} + O\left(\frac{a^4}{n^3} + \frac{1}{a}\right)\right).$$

For $(x, y) \in \mathbb{N}^2$, we have

$$(7) \quad \int_0^1 t^x (1-t)^y dt = \frac{x!}{y!} (x+y+1)! = \frac{1}{(x+y+1) \binom{x+y}{x}}.$$

Setting $N = \binom{n}{b} - \binom{n-a}{b}$, using standard calculus we then obtain

$$(8) \quad N = \frac{n^{(b-1)a}}{(b-1)!} \left(1 - \frac{a(b-1)}{2n} + \frac{a^2(b-1)(b-2)}{6n^2} + O\left(\frac{b}{n}\right) + O\left(\frac{ab^3}{n^2} + \frac{b^4}{n^2}\right)\right).$$

Now, using the above formulas we easily find that the integral equals

$$(9) \quad \begin{aligned} \frac{1}{N \binom{N}{k-1}} &= \frac{\sqrt{2\pi k}}{N^k} \frac{(k-1)^{(k-1)}}{e^{k-1}} \left(1 + O\left(\frac{k^2}{N} + \frac{1}{k}\right)\right) \\ &= \sqrt{\frac{2\pi}{k}} \frac{k^k}{N^k e^k} \left(1 + O\left(\frac{k^2}{N} + \frac{1}{k}\right)\right) \\ &= \sqrt{\frac{2\pi}{k}} \frac{k^k}{e^k} \frac{[(b-1)!]^k}{n^{k(b-1)} a^k} \left(1 + O\left(\frac{k}{n^{b-1}b} + \frac{1}{k}\right)\right) \\ &\times \exp\left(-k \log\left(1 - \frac{a(b-1)}{2n} + \frac{a^2(b-1)(b-2)}{6n^2} + O\left(\frac{b}{n}\right) + O\left(\frac{ab^3}{n^2} + \frac{b^4}{n^2}\right)\right)\right). \end{aligned}$$

Therefore by replacing a with $k(b-1) - \ell$ and using (6), it yields (8)

$$\frac{\binom{n}{a}}{N \binom{N}{k-1}} \sim \frac{1}{\sqrt{(b-1)} n^\ell} \frac{k^{(k-1)} [(b-1)!]^k \exp(k(b-2) - \ell) \exp\left(-\frac{(b-1)^4 k^3}{24 n^2}\right)}{\left(k(b-1) - \ell\right)^{kb-\ell}},$$

where \sim means that the asymptotic equation holds up to a factor of

$$1 + O\left(\frac{bk}{n} + \frac{b^4 k^2}{n^2} + \frac{b^4 k}{n^2} + \frac{b^4 k^4}{n^3} + \frac{k}{n^{b-1}b} + \frac{1}{k}\right).$$

□

Lemma 2.2 tells us that the expectations the random variables of interest rely on the asymptotic number of the considered connected components.

2.2. Asymptotic enumeration of connected hypergraphs. As far as we know there are not so many results about the asymptotic enumeration of connected uniform hypergraphs. In this paragraph, we recall some of the results established independently in [20, 7, 1] (the three papers actually use three different methods). In [1], the authors use the generating functions approach [15, 18, 11, 29, 30, 31] to count exactly and asymptotically connected labeled b -uniform hypergraphs. If $A(z) = \sum_n a_n z^n$ and $B(z) = \sum_n b_n z^n$ are two formal power series, $A \preceq B$ means that $\forall n \in \mathbb{N}$, $a_n \leq b_n$. Among other results, the authors of [1] established the following:

Lemma 2.3. *Let $H_\ell(z)$ be the exponential generating function (EGF for short) of b -uniform connected hypergraphs with excess ℓ . Define by $T(z)$ the EGF of labeled rooted hypertrees. Then,*

$$(10) \quad H_{-1}(z) = T(z) - \frac{(b-1)T(z)^b}{b!} \quad \text{with} \quad T(z) = z \exp\left(\frac{T(z)^{(b-1)}}{(b-1)!}\right) = z \frac{\partial H_{-1}(z)}{\partial z}.$$

For any $\ell \geq 1$, H_ℓ satisfies

$$(11) \quad \frac{\lambda_\ell(b-1)^{2\ell}}{3\ell T(z)^\ell \theta(z)^{3\ell}} - \frac{(\nu_\ell(b-2))(b-1)^{2\ell-1}}{(3\ell-1)T(z)^\ell \theta(z)^{3\ell-1}} \preceq H_\ell(z) \preceq \frac{\lambda_\ell(b-1)^{2\ell}}{3\ell T(z)^\ell \theta(z)^{3\ell}},$$

where $\lambda_\ell = 3 \left(\frac{3}{2}\right)^\ell \frac{\ell!}{2\pi} (1 + O(\frac{1}{\ell}))$ and $\nu_\ell = O(\ell\lambda_\ell)$. Furthermore, λ_ℓ is defined recursively by $\lambda_0 = \frac{1}{2}$ and

$$(12) \quad \lambda_\ell = \frac{1}{2}\lambda_{\ell-1}(3\ell-1) + \frac{1}{2}\sum_{t=0}^{\ell-1}\lambda_t\lambda_{\ell-1-t}, \quad (\ell \geq 1).$$

We also need the following result which has been proved independently by Karoński and Łuczak in [20] and Andriamampianiana and Ravelomanana in [1]:

Lemma 2.4. *For $\ell \equiv \ell(n)$ such that $\ell = o(\sqrt[3]{\frac{n}{b}})$ as $n \rightarrow \infty$, the number of connected b -uniform hypergraphs built with n vertices and having excess ℓ satisfies*

$$(13) \quad \sqrt{\frac{3}{2\pi}} \frac{(b-1)^{\frac{\ell}{2}} e^{\frac{\ell}{2}} n^{n+\frac{3\ell}{2}-\frac{1}{2}}}{12^{\frac{\ell}{2}} \ell^{\frac{\ell}{2}} ((b-2)!)^{\frac{n+\ell}{b-1}}} \exp\left(\frac{n}{b-1} - n\right) \left(1 + O\left(\frac{1}{\sqrt{\ell}}\right) + O\left(\sqrt{\frac{b\ell^3}{n}}\right)\right).$$

Observe that setting $b = 2$ in (13), we retrieve the asymptotical results of Sir E. M. Wright for connected graphs in his fundamental paper [31].

3. HYPERTREES AND UNICYCLIC COMPONENTS

As typical examples, let us work with unicyclic components. We will compute the expected number of transitions $-1 \rightarrow 0$. That is the number of times unicyclic connected components (i.e. 0-components) are created. We will also investigate the number of times unicyclic components merge with hypertrees growing in size but staying with the same complexity (excess 0). In these directions, we have the following result :

Theorem 3.1. *As $n \rightarrow \infty$, on the average a b -uniform random hypergraph has about $\frac{1}{3} \log n$ dynamic unicyclic components. The expected number of static 0-components is $\sim \frac{\sqrt{2}\pi^{3/2}24^{1/6}}{6\Gamma(\frac{5}{6})} (b-1)^{1/3} n^{1/3} \approx 1.975 (b-1)^{1/3} n^{1/3}$.*

Proof. The creation of unicyclic components can be obtained only by adding an edge joining 2 distinct vertices inside the same hypertree with $(b-2)$ other vertices from $(b-2)$ distinct hypertrees (to complete the edge).

The number of such constructions is therefore given by the coefficients of the following EGF :

$$(14) \quad C'_0(z) = \frac{(\vartheta_z H_{-1}(z))^{(b-2)}}{(b-2)!} \times \left(\frac{\vartheta_z^2 - \vartheta_z}{2} (H_{-1}(z)) \right),$$

where the combinatorial operator $\vartheta_z = z \frac{\partial}{\partial z}$ corresponds to marking a vertice of the hypergraph in order to distinguish it from the others. For instance, we refer the reader to Bergeron, Labelle and Leroux [3] for the use of distinguishing/markings and pointing in combinatorial species. Recall that the EGFs are as described briefly in Lemma 2.3 (see also the Appendix), using $\vartheta_z H_{-1}(z) = T(z)$ and $\vartheta_z T(z) = \frac{T(z)}{1 - T(z)^{(b-1)}/(b-2)!}$ we find

$$(15) \quad C'_0(z) = \frac{T(z)^{b-2}}{2(b-2)!} \left(\frac{T(z)}{1 - \frac{T(z)^{(b-1)}}{(b-2)!}} - T(z) \right) = \frac{1}{2} \left(\frac{1}{1 - \frac{T(z)^{(b-1)}}{(b-2)!}} - T(z) - 1 \right).$$

We also have (such expansions are similar to those in [22])

$$(16) \quad \frac{1}{1 - \frac{T(z)^{(b-1)}}{(b-2)!}} = \sum_{k=0}^{\infty} \frac{k^k}{k! [(b-2)!]^k} z^{(b-1)k}.$$

Denoting by $\rho'((b-1)k, k)$ the number of manners to label a unicyclic component with $(b-1)k$ vertices and with a distinguished edge such that its deletion will leave a forest of hypertrees, we thus have

$$(17) \quad \rho'((b-1)k, k) = ((b-1)k)! [z^{(b-1)k}] C'_0(z) \sim \frac{((b-1)k)! k^k}{2 k! [(b-2)!]^k}$$

(where if $A(z) = \sum_n a_n z^n$ then $[z^n] A(z) = a_n$).

Next, using Lemma 2.2 with the above equation, after nice cancellations and summing over all possible values of k , we get

$$(18) \quad \sum_{k=1}^{\frac{n}{b-1}} \rho'((b-1)k, k) \binom{n}{(b-1)k} \int_0^1 t^{k-1} (1-t)^{\binom{n}{b} - \binom{n-(b-1)k}{b} - k} dt \\ \sim \sum_{k=1}^{\frac{n}{b-1}} \frac{1}{2k} \times \exp\left(-\frac{(b-1)^4 k^3}{24n^2}\right) \sim \frac{1}{2} \int_1^{n/(b-1)} \frac{1}{x} e^{-(b-1)^4 x^3/24n^2} dx.$$

To estimate the last integral, we write

$$(19) \quad \int_1^{n/(b-1)} \frac{1}{x} e^{-(b-1)^4 x^3/24n^2} dx = \int_1^{n^{2/3}/(b-1)^{4/3}} \frac{1}{x} \left(1 + O\left(\frac{(b-1)^4 x^2}{n^2}\right)\right) dx \\ + O\left(\int_{n^{2/3}/(b-1)^{4/3}}^{n/(b-1)} \frac{1}{x} e^{-(b-1)^4 x^3/24n^2} dx\right) \sim \log n^{2/3} + O(1).$$

Thus, the expected number of creations of unicyclic components is $\sim \frac{1}{3} \log n$. which completes the proof of the first part of the theorem. To prove the second part, we have to investigate the number of static 0-components, that is the number of times 0-components merge with hypertrees by the transition $0 \rightarrow 0$. The EGF of unicyclic components with a distinguished edge such that its suppression will leave a unicyclic component and a set of $(b-2)$ rooted hypertrees is given by

$$(20) \quad C''_0(z) = \frac{T(z)^{b-2}}{(b-2)!} \vartheta_z(H_0(z)) = \frac{T(z)^{b-2}}{(b-2)!} \left(\frac{(b-1)T(z)^{b-1}}{2(b-2)!\theta^2} - \frac{(b-1)T(z)^{b-1}}{2(b-2)!\theta} \right)$$

where $\theta = 1 - T(z)^{b-1}/(b-2)!$. Denote by $\rho''((b-1)k, k)$ the number of manners to label a unicyclic component with $(b-1)k$ vertices and with a distinguished edge such that its deletion will leave a 0-component with a forest of rooted hypertrees, we claim that

$$(21) \quad \rho''((b-1)k, k) = ((b-1)k)! \left[z^{(b-1)k} \right] C''_0(z) \sim \sqrt{\frac{\pi(b-1)^3}{8}} \frac{k^{k(b-1)+1/2}}{e^{k(b-2)}} \left(\frac{(b-1)^{k(b-1)}}{[(b-2)!]^k} \right).$$

(We omit the details, since the full proof involves singularity analysis [11] of the EGF C''_0 described above.)

Now, using Lemma 2.2 and summing over k after some cancellations, the computed expectation is about

$$(22) \quad \sum_{k=1}^{n/(b-1)} \sqrt{\frac{\pi}{8}} (b-1) \frac{1}{k^{1/2}} e^{-(b-1)^4 k^3/24n^2} \sim \sqrt{\frac{\pi}{8}} (b-1) \int_1^{n/(b-1)} \frac{e^{-(b-1)^4 x^3/24n^2}}{\sqrt{x}} dx \\ \sim 1/6 \frac{\sqrt{2\pi^{3/2} 24^{1/6}}}{\Gamma(5/6)} (b-1)^{1/3} n^{1/3} \approx 1.974748319 \dots (b-1)^{1/3} n^{1/3}.$$

□

Note here that the result stated in Theorem 3.1 (humbly) generalizes the ones of Janson in [17] since by setting $b = 2$, we retrieve his results concerning unicyclic (graph) components.

Next, we can investigate the number of vertices that ever belong to 0-components. According to the above computations, the expected number of vertices added to \mathcal{V}_0 for the

creation of such unicyclic components (transition $-1 \rightarrow 0$) is about

$$(23) \quad \frac{1}{2} \sum_{k=1}^{n/(b-1)} (b-1) e^{-(b-1)^4 k^3 / 24n^2} \sim \frac{1}{6} \frac{24^{1/3} \Gamma(1/3)}{(b-1)^{1/3}} n^{2/3}.$$

Whenever the excess ℓ is fixed, that is $\ell = O(1)$, the methods developped here for unicyclic components can be generalized, using analytical tools such those in [11]. Thus, we now turn on components with higher complexities.

4. MULTICYCLIC COMPONENTS WITH EXTREMAL COMPLEXITIES

In this section, we focus on the creation and growth of components of higher complexity. First, we will compute the expectations of the number of creations of ℓ -components for $\ell \geq 1$. To this purpose, we need several intermediate lemmas.

Define $h_n(\xi, \beta)$ as follows

$$(24) \quad \frac{1}{T(z)^\xi \left(1 - \frac{T(z)^{b-1}}{(b-2)!}\right)^{3\xi+\beta}} = \sum_{n \geq 0} h_n(\xi, \beta) \frac{z^n}{n!}.$$

The following lemma is an application of the saddle point method [5, 11] which is well suited to cope with our analysis :

Lemma 4.1. *Let $\xi \equiv \xi(n)$ be such that $\xi(b-1) \rightarrow 0$ but $\frac{\xi(b-1)n}{\ln n^2} \rightarrow \infty$ and let β be a fixed number. Then $h_n(\xi n, \beta)$ defined in (24) satisfies*

$$(25) \quad \begin{aligned} h_n(\xi n, \beta) &= \frac{n!}{\sqrt{2\pi n(b-1)} \left((b-1)!\right)^{\frac{\xi n+n}{b-1}}} \left(1 - (b-1)u_0\right)^{(1-\beta)} \\ &\times \exp(n\Phi(u_0)) \left(1 + O\left(\sqrt{\xi(b-1)}\right) + O\left(\frac{1}{\sqrt{\xi(b-1)n}}\right)\right), \end{aligned}$$

where

$$(26) \quad \begin{aligned} \Phi(u) &= u - \left(\frac{\xi+1}{b-1}\right) \ln u - 3\xi \ln(1 - (b-1)u) \\ u_0 &= \frac{3/2\xi b - \xi + 1 - 1/2\sqrt{\Delta}}{b-1} \quad \text{with } \Delta = 9\xi^2 b^2 - 12\xi^2 b + 12\xi b + 4\xi^2 - 12\xi. \end{aligned}$$

Proof. One can start with Cauchy's integral formula. Note that the radius of convergence of the series $T(z)$ is given by $\sqrt[b-1]{(b-2)!} \exp(-1/(b-1))$. We make the substitution $u = T(z)^{(b-1)}/(b-1)!$ and get successively

$$(27) \quad \begin{aligned} T(z) &= \sqrt[b-1]{(b-1)!} u, \quad z = \sqrt[b-1]{(b-1)!} u e^{-u} \quad \text{and} \\ dz &= \left(\frac{1}{(b-1)u} - 1\right) \left((b-1)!\right)^{\frac{1}{b-1}} e^{-u} du. \end{aligned}$$

From the Cauchy integral formula, we then obtain

$$(28) \quad h_n(\xi n, \beta) = \frac{n!}{2\pi i \left((b-1)!\right)^{(\xi n+n)/(b-1)}} \oint \frac{(1 - (b-1)u)^{1-\beta}}{(b-1)u} \exp(n\Phi(u)) du,$$

where $\Phi(u) = u - \left(\frac{\xi+1}{b-1}\right) \ln u - 3\xi \ln(1 - (b-1)u)$. The big power in the integrand, viz. $\exp(n\Phi(u))$, suggests us to use the saddle point method. Investigating the roots of $\Phi'(u) = 0$, we find two saddle points, $u_0 = \frac{3/2\xi b - \xi + 1 - 1/2\sqrt{\Delta}}{b-1}$ and $u_1 = \frac{3/2\xi b - \xi + 1 + 1/2\sqrt{\Delta}}{b-1}$ with $\Delta = 9\xi^2 b^2 - 12\xi^2 b + 12\xi b + 4\xi^2 - 12\xi$. Moreover, we have $\Phi''(u) = \frac{\xi+1}{(b-1)u^2} + 3\frac{\xi(-b+1)^2}{(1-(b-1)u)^2}$ so that for $u \notin \{0, 1/(b-1)\}$, $\Phi''(u) > 0$. The main point of the application of the saddle point method here is that $\Phi'(u_0) = 0$ and $\Phi''(u_0) > 0$, hence $n\Phi(u_0 \exp(i\tau))$ is well approximated by $n\Phi(u_0) - nu_0^2 \Phi''(u_0) \frac{\tau^2}{2}$ in the vicinity of $\tau = 0$. If we integrate (28) around a circle passing vertically through $u = u_0$ in the z -plane, we obtain

$$(29) \quad h_n(\xi n, \beta) = \frac{n!}{2\pi \left((b-1)!\right)^{(\xi n + n)/(b-1)}} \int_{-\pi}^{\pi} \frac{(1 - (b-1)u_0 e^{i\tau})^{1-\beta}}{(b-1)} \exp(n\Phi(u_0 e^{i\tau})) d\tau$$

where

$$(30) \quad \Phi(u_0 e^{i\tau}) = u_0 \cos \tau + iu_0 \sin \tau - \frac{\xi+1}{b-1} \ln u_0 - i\frac{\xi+1}{b-1} \tau - 3\xi \ln(1 - (b-1)u_0 e^{i\tau}).$$

Denoting by $\Re(z)$ the real part of z , if $f(\tau) = \Re(\Phi(u_0 e^{i\tau}))$ we have

$$(31) \quad f(\tau) = u_0 \cos \tau - \frac{\xi+1}{b-1} \ln u_0 - 3\xi \ln u_0 - 3\xi \ln(b-1) - \frac{3\xi}{2} \ln \left(1 + \frac{1}{(b-1)^2 u_0^2} - \frac{2 \cos \tau}{(b-1)u_0}\right).$$

It comes

$$(32) \quad f'(\tau) = \frac{d}{d\tau} \Re(h(u_0 e^{i\tau})) = -u_0 \sin \tau - \frac{3\xi \sin \tau}{u_0(b-1) + \frac{1}{(b-1)u_0} - 2 \cos \tau}.$$

Therefore, if $\tau = 0$ $f'(\tau) = 0$. Also, $f(\tau)$ is a symmetric function of τ and in $[-\pi, -\tau_0] \cup [\tau_0, \pi]$, for any given $\tau_0 \in (0, \pi)$, and $f(\tau)$ takes its maximum value for $\tau = \tau_0$. Since $|\exp(\Phi(u))| = \exp(\Re(\Phi(u)))$, when splitting the integral in (29) into three parts, viz. " $\int_{-\pi}^{-\tau_0} + \int_{-\tau_0}^{\tau_0} + \int_{\tau_0}^{\pi}$ ", we know that it suffices to integrate from $-\tau_0$ to τ_0 , for a convenient value of τ_0 , because the others can be bounded by the magnitude of the integrand at τ_0 . In fact, we have $\Phi(u_0 e^{i\theta}) = \Phi(u_0) + \sum_{p \geq 2} \phi_p (e^{i\theta} - 1)^p$ with $\phi_p = \frac{u_0^p}{p!} \Phi^{(p)}(u_0)$. We easily compute $\Phi^{(p)}(u_0) = (-1)^p (p-1)! \left(\frac{\xi+1}{(b-1)u_0^p} + \frac{3\xi(1-b)^p}{(1-(b-1)u_0)^p} \right)$, for $p \geq 2$. Whenever $\xi b \rightarrow 0$, we have $(b-1)u_0 = 1 - \sqrt{3(b-1)\xi} + (3/2b-1)\xi + O(b^{3/2}\xi^{3/2})$. Therefore, we obtain after a bit of algebra

$$(33) \quad |\phi_p| \leq O\left(\frac{2^p}{\xi^{\frac{p}{2}-1}(b-1)^{\frac{p}{2}}}\right), \quad \text{as } \xi(b-1) \rightarrow 0.$$

On the other hand,

$$(34) \quad |e^{i\tau} - 1| = \sqrt{2(1 - \cos \tau)} < \tau, \quad \tau > 0.$$

Thus, the summation can be bounded for values of τ and ξ such that $\tau \rightarrow 0$, $\xi b \rightarrow 0$ ($\xi \rightarrow 0$) but $\frac{\tau}{\sqrt{\xi}} \rightarrow 0$ and we have

$$(35) \quad \left| \sum_{p \geq 4} \phi_p (e^{i\tau} - 1)^p \right| \leq \sum_{p \geq 4} |\phi_p \tau^p| \leq \sum_{p \geq 4} O\left(\frac{2^p \tau^p}{\xi^{\frac{p}{2}-1}(b-1)^{\frac{p}{2}}}\right) = O\left(\frac{\tau^4}{\xi(b-1)}\right).$$

It follows that for $\tau \rightarrow 0$, $\xi(b-1) \rightarrow 0$ and $\frac{\tau}{\sqrt{\xi(b-1)}} \rightarrow 0$, $\Phi(u_0 e^{i\tau})$ can be rewritten as

$$(36) \quad \begin{aligned} \Phi(u_0 e^{i\tau}) &= \Phi(u_0) - \frac{1}{(b-1)} \left(1 - \frac{\sqrt{\xi}}{\sqrt{3(b-1)}} \frac{3b-4}{2} + \frac{(9b^2-12b+4)}{12(b-1)} \xi \right) \tau^2 \\ &- \frac{i}{(b-1)} \left(1 - \frac{(3b-4)\sqrt{\xi}}{2\sqrt{3(b-1)}} + \frac{(9b^2-12b+4)}{12(b-1)} \xi \right) \tau^3 + O\left(\frac{\tau^4}{\xi(b-1)}\right). \end{aligned}$$

Therefore, if $\xi(b-1) \rightarrow 0$ but $\frac{\xi(b-1)n}{(\ln n)^2} \rightarrow \infty$, if we let $\tau_0 = \frac{\ln n}{\sqrt{n u_0^2 \Phi''(u_0)}}$ (with $u_0^2 \Phi''(u_0) = \frac{2}{b-1} + O(\sqrt{\xi(b-1)})$) we can remark (as already said) that it suffices to integrate (29) from $-\tau_0$ to τ_0 , using the magnitude of the integrand at τ_0 to bound the resulting error. The rest of the proof is now standard application of the saddle point method (see for instance De Bruijn [5, Chapters 5 & 6]) and is omitted in this extended abstract. After a bit of algebra, one gets the formula (25). \square

Lemma 4.2. *Let $a = k(b-1) - \ell$. Denote by $c_\ell(a, k)$ the number of manners to label an ℓ -component with a vertices such that one edge – whose deletion will suppress the occurrence of the created ℓ -component – is distinguished among the others. As ℓ tends to ∞ with the number of vertices a such that $\ell = o(\sqrt[3]{\frac{a}{b}})$ then*

$$(37) \quad c_\ell(a, k) = a! [z^a] \left(\frac{1}{2} \frac{(3\ell)(b-1)^{2\ell} \lambda_{\ell-1}}{T(z)^\ell \theta^{3\ell+1}} \right) \times \left(1 + O\left(\frac{1}{\sqrt{\ell}}\right) + O\left(\sqrt{\frac{b\ell^3}{a}}\right) \right),$$

where $\theta = 1 - T(z)^{b-1}/(b-2)!$ and the sequence (λ_ℓ) is defined with (12).

Sketch of proof. The proof given in this extended abstract is sketched. The main ideas are as follows. The inequalities given by equation (11) in Lemma 2.3 tell us that when ℓ is large, the main constructions that lead to the creation of a new ℓ -component arises a.a.s. from picking two distinct vertices in an $(\ell-1)$ -component and joining them by an edge with $(b-2)$ set of rooted hypertrees. Such constructions are counted by

$$\left(\frac{\vartheta_z^2 - \vartheta_z}{2} H_{\ell-1}(z) \right) \times \frac{T(z)^{b-2}}{(b-2)!}.$$

Using again (11) with (25), one can show that the coefficient of the latter EGF has the same asymptotical behaviour as the following one

$$\frac{3\ell(b-1)^{2\ell} \lambda_{\ell-1}}{2 T(z)^\ell \theta^{3\ell+1}}.$$

(The error terms being the same as those given by the saddle-point Lemma 4.1 above.) \square

We then have the following result :

Theorem 4.3. *As $\ell, b \rightarrow \infty$ with n but such that $\ell = o(\sqrt[3]{\frac{n}{b}})$, the expected number of creations of ℓ -component is ~ 1 and the expected number of vertices that ever belong to an ℓ -component is about $(12\ell(b-1))^{1/3} n^{2/3}$. Thus, $\mathbb{E}[V_\ell^{max}] = O(\ell^{1/3}(b-1)^{1/3} n^{2/3})$.*

Sketch of the proof. The proof of this Theorem is a combination of Lemmas 4.1, 4.2 and 2.2 together with the asymptotic value of λ_ℓ given in Lemma 2.3 and with the fact that

$$\sum_{k=0}^{n/(b-1)} k^u \exp\left(-\frac{(b-1)^4 k^3}{24 n^2}\right) \sim \frac{1}{3} \frac{24^{\frac{u+1}{3}} n^{\frac{2(u+1)}{3}}}{(b-1)^{\frac{4(u+1)}{3}}} \Gamma\left(\frac{u+1}{3}\right).$$

REFERENCES

- [1] Andriamampianina T. and Ravelomanana V. (2005). Enumeration of connected uniform hypergraphs. *In Proc. of the 17-th Formal Power Series and Algebraic Combinatorics.*
- [2] Berge, C. (1976). *Graphs and hypergraphs* North-Holland Mathematical Library.
- [3] Bergeron, F., Labelle, G. and Leroux, P. (1998). *Combinatorial Species and Tree-like Structures* Encyclopedia of Mathematics and its Applications, Vol. 67, Cambridge Univ. Press.
- [4] Bollobás, B. (1985) *Random Graphs* Academic Press, London.
- [5] De Bruijn, N. G. (1981). *Asymptotic Methods in Analysis.* Dover, New-York.
- [6] Cayley, A. (1889). A Theorem on Trees. *Quart. J. Math. Oxford Ser.*, 23:376–378.
- [7] Coja-Oghlan, A., Moore, C. and Sanwalani, V. (2004). Counting Connected Graphs and Hypergraphs via the Probabilistic Method. In Proc. of APPROX-RANDOM 2004.
- [8] Erdős, P. and Rényi A. (1959). On random graphs. *Publ. Math. Debrecen*, 6:290–297.
- [9] Erdős, P. and Rényi A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5:17–61.
- [10] Flajolet, P., Knuth, D. E. and Pittel B. (1989). The first cycles in an evolving graph. *Discrete Math.*, 75:167–215.
- [11] Flajolet, P. and Sedgewick, R. *Analytic Combinatorics.* To appear (chapters are available as Inria research reports). See <http://algo.inria.fr/flajolet/Publications/books.html>.
- [12] Flajolet, P. and Sedgewick, R. *The average case analysis of algorithms: Counting and generating functions*
- [13] Goulden, I. P. and Jackson, D. M. (1983). *Combinatorial Enumeration.* Wiley, New York.
- [14] Harary, F. and Palmer, E. (1973). *Graphical Enumeration.* Academic Press, New-York and London.
- [15] Herbert S. Wilf *Generatingfunctionology*
- [16] Janson, S. (1993). Multicyclic components in random graphs process. *Random Structures and Algorithms*, 4:71–84.
- [17] Janson, S. (2000). Growth of components in random graphs. *Random Structures and Algorithms*, 17:343–356.
- [18] Janson, S., Knuth, D. E., Łuczak, T. and Pittel B. (1993). The birth of the giant component. *Random Structures and Algorithms*, 4:233–358.
- [19] Janson, S., Łuczak, T. and Ruciński A. (2000). *Random Graphs.* John Wiley, New York.
- [20] Karoński, M. and Łuczak, T. (1997). The number of sparsely edges connected uniform hypergraphs. *Discrete Math.* 171:153–168.
- [21] Karoński, M. and Łuczak, T. and (2002). The phase transition in a random hypergraph. *J. Comput. Appl. Math.* 142:125–135.
- [22] Knuth, D. E. and Pittel, B. (1989). A recurrence related to trees. *Proc. Am. Math. Soc.*, 105:335–349.
- [23] Pittel, B. and Wormald N. (2003). Counting connected graphs inside-out. *J. Comb. Th. Series B*, to appear.
- [24] Ravelomanana, V. (to appear). The average size of giant components between the double-jump. *Algoritmica*.
- [25] Rényi, A. (1959). On connected graphs I. *Publ. Math. Inst. Hungarian Acad. Sci.* 4:385–388.
- [26] Schmidt-Pruzan, J. and Shamir S. (1985). Component structure in the evolution of random hypergraphs *Combinatorica* 5: 81–94.
- [27] Selivanov, B. I. (1972). Perechislenie odnorodnykh hipergrafov c protstoi ciklovoi strukturoi, Kombinatornyi Analiz 2, pages 60–67.
- [28] Voblyi, V. A. O koeffitsientakh Raıta i Stepanova-Raıta *Matematicheskie Zametki* 42 (1987), 854–862. English translation, V. A. Voblyi, “Wright and Stepanov-Wright coefficients,” *Mathematical Notes* 42 (1987), 969–974.
- [29] Wright, E. M. (1977). The Number of Connected Sparsely Edged Graphs. *Journal of Graph Theory*, 1:317–330.
- [30] Wright, E. M. (1978). The Number of Connected Sparsely Edged Graphs. II. Smooth graphs and blocks *Journal of Graph Theory*, 2:299–305.
- [31] Wright, E. M. (1980). The Number of Connected Sparsely Edged Graphs. III. Asymptotic results *Journal of Graph Theory*, 4:393–407.

VLADY RAVELOMANANA (CORRESPONDING AUTHOR), LIPN, UNIVERSITÉ DE PARIS-NORD, F 93430 VILLETANEUSE, FRANCE.

ALPHONSE LAZA RIJAMAMY, UNIVERSITÉ DE MADAGASCAR, ANKATSO – TANA 101, MADAGASCAR.